

**100% ACTUAIRES &  
100% DATA SCIENCE**

INSTITUT DES  
**ACTUAIRES**



**29 / NOV / 2019**

Hôtel Marriott Rive Gauche  
Paris 14ème

# **Risque Cyber: comment construire des modèles de tarification, de provisionnement, et d'accumulation à partir de données extrêmes, hétérogènes et surtout peu nombreuses ?**

- Introduction: enjeux actuariels pour une compagnie d'assurance
- Présentation de la base Privacy Rights Clearinghouse (PRC)
- Analyse de la sévérité: distinguer les différents comportements
- Etude de la fréquence: calibrer le phénomène d'auto excitation
- Comprendre la diffusion des événements Cyber et adapter la réponse
- Conclusion

# Introduction aux enjeux actuariels du Cyber risque pour une compagnie d'assurance

## Deux types de risques Cyber sont supportés par les assureurs :

- le risque Cyber propre (Operational Risk) ;
- le risque Cyber assurantiel (P&C Risk).

Dans cet atelier nous nous intéressons surtout au second. A la différence du premier, nous avons une connaissance moins profonde du risque et devons avoir recours à des approches plus statistiques.

## Les assureurs couvrent le risque Cyber sous diverses formes...

- Produit Cyber : contrat disposant de clauses spécifiques à l'assurance Cyber
- Péril Cyber : contrat traditionnel de type Dommage ou RC dans lequel le péril Cyber a explicitement été inclus comme déclencheur potentiel

### ... et parfois sans même le savoir

- « Silent » Cyber : le péril Cyber n'est pas toujours explicitement inclus ou exclu des contrats d'assurance traditionnel, dans ce cas on parle de « Silent » Cyber

D'un point de vue modélisation, le Cyber est embarqué dans une structure similaire à ce qui existe sur les autres LoB P&C :

- risque attritionnel/atypique : sinistres individuels de type perte d'ordinateur, arnaque au dirigeant, phishing ciblé, etc.
- risque d'accumulation: équivalent du risque Catastrophe Naturel : attaque majeure affectant un grand nombre d'assurés (caractère systémique) au sein d'un même évènement.

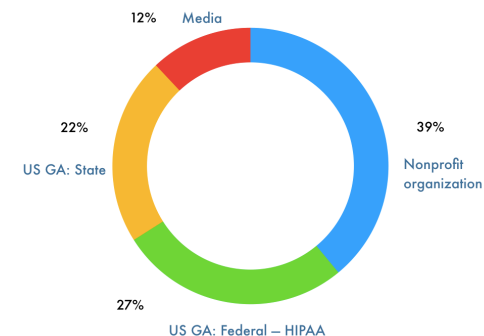
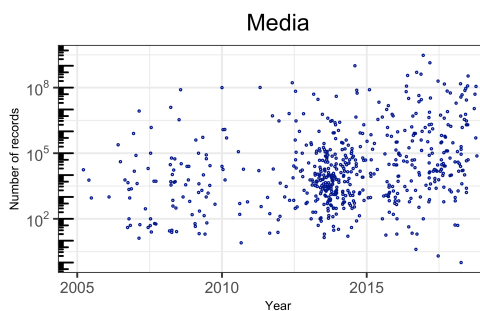
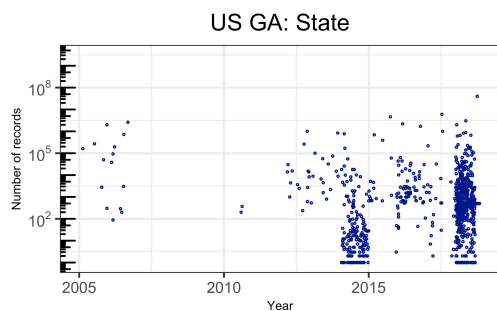
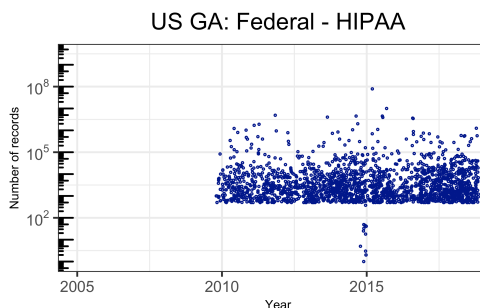
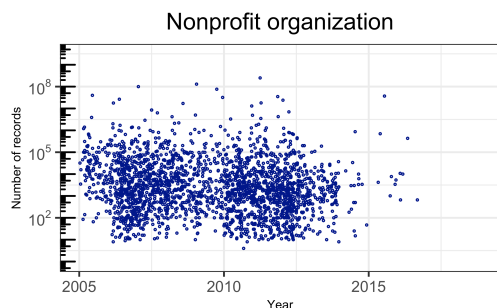
La faible maturité du risque incite aux approches de type **scénario** dans lesquels nous pouvons imaginer des évènements Cyber sensiblement plus graves que ceux observés à l'heure actuelle.

Toutefois, l'existence de bases de données externes permet aussi des approches de type **statistique**.

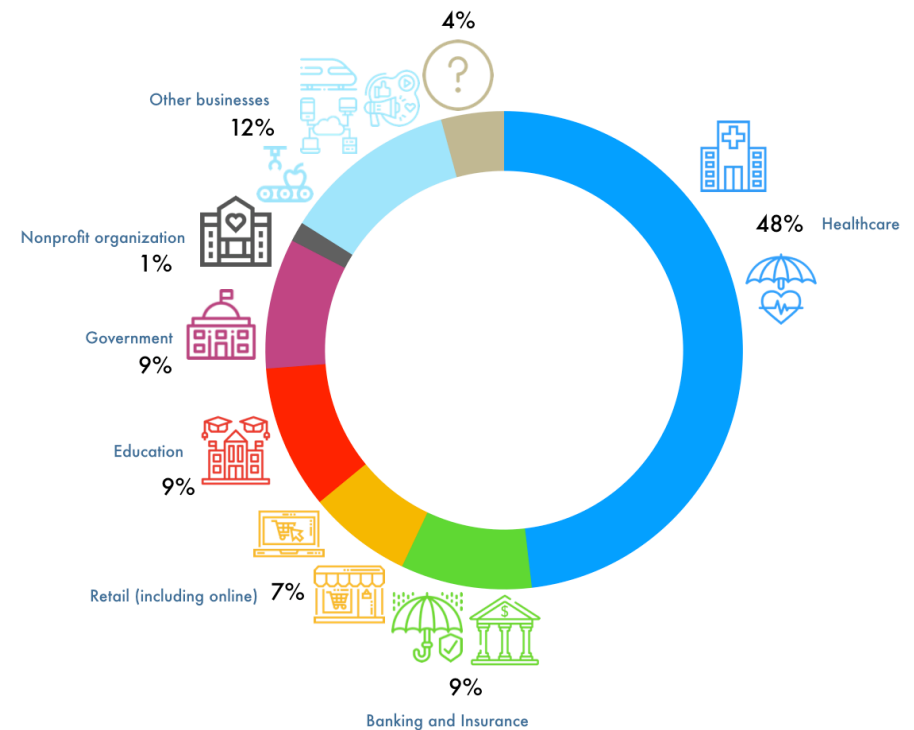
# Présentation de la base Privacy Rights Clearinghouse (PRC)



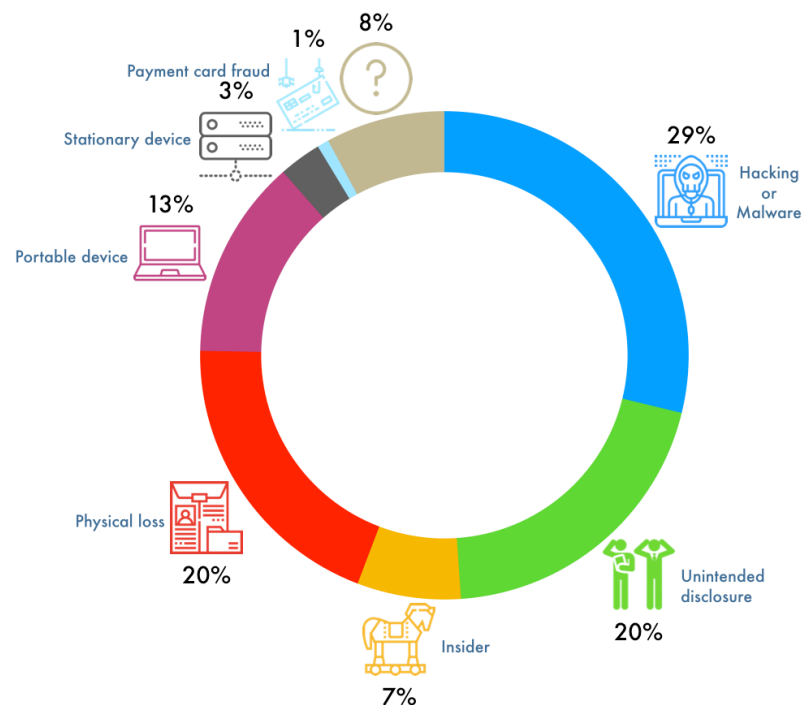
- Privacy Rights Clearinghouse
- Plus de 8000 failles de données référencées depuis 2005
- Événements listés grâce au regroupement de différentes sources:



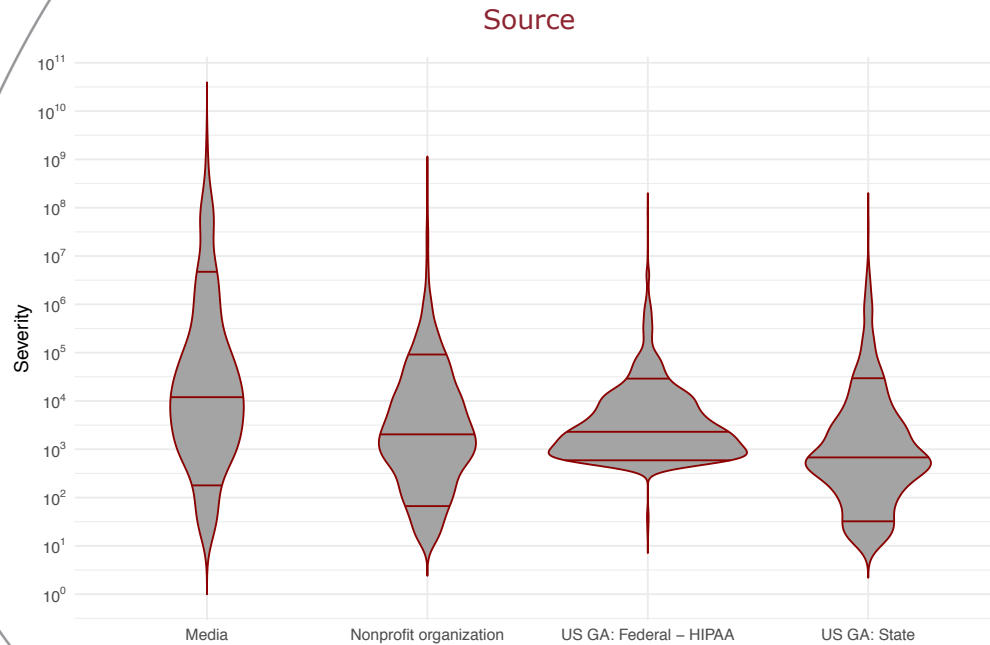
- Informations concernant l'entreprise (exposition):
  - le nom,
  - le secteur d'activité,
  - la localisation.



- Informations à propos de l'événement:
  - le canal de notification,
  - la date de notification,
  - l'origine de la faille de données,
  - le nombre de ligne,
  - une description de l'événement.

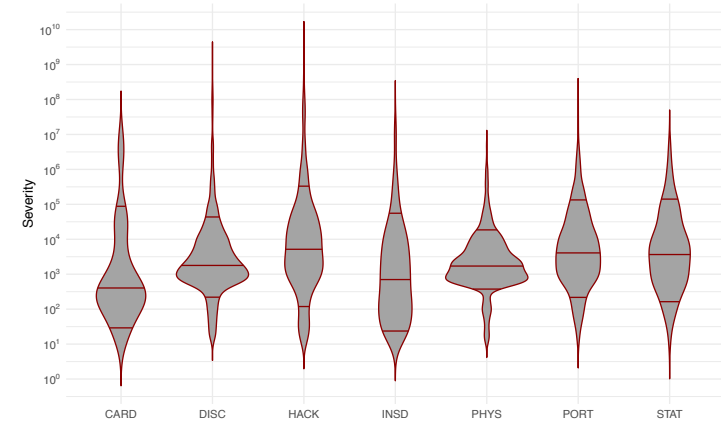


## Hétérogénéité de la sévérité

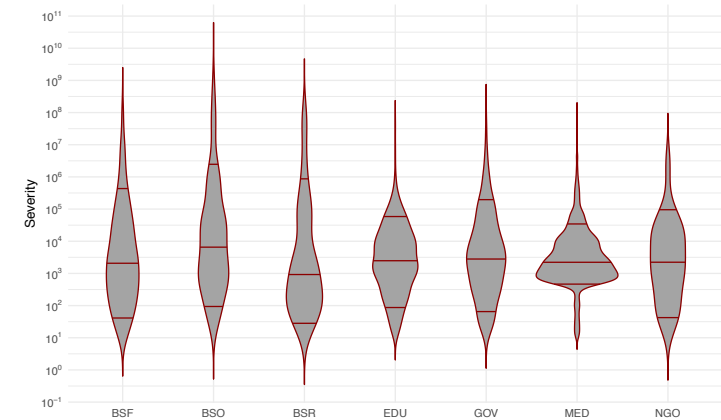


Médiane: 2 000  
Moyenne: 1 821 682

## Type d'événement



## Secteur d'activité



# Analyse de la sévérité: distinguer les différents comportements

en collaboration avec  
Olivier Lopez (Sorbonne Université)  
et  
Maud Thomas (Sorbonne Université)



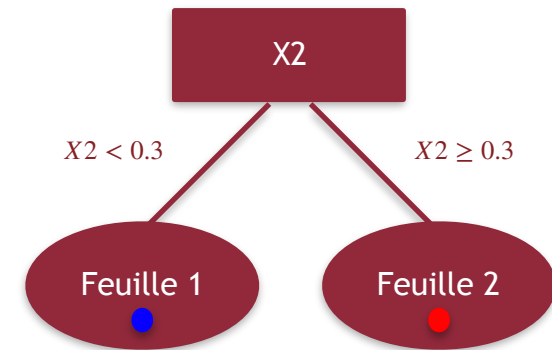
- Etude de la sévérité des événements Cyber
- Facteurs de risque:
  - identification,
  - impacts sur un scénario central,
  - impacts sur un scénario extrême.
- Applications:
  - assurabilité,
  - souscription et couverture contractuelle,
  - provisionnement et appétit au risque.

- Algorithme CART, Breiman (1984)

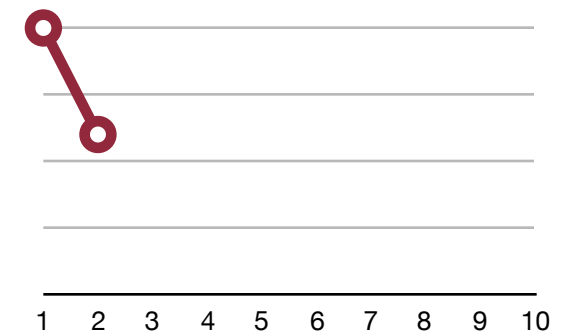
- Objectif:

$$m^* = \arg \min_{m \in \mathcal{M}} E[\phi(Y, m(\mathbf{X}))]$$

- $Y$  variable réponse
- $\mathbf{X} \in \mathcal{X} \subset \mathbb{R}^d$  variables explicatives



Réduction de l'erreur



## Des fonctions de pertes à choisir en fonction de l'objectif

Objectifs:

- Espérance:

$$\mathbb{E}[Y|X]$$

- Médiane:

$$F_{Y|X}^{-1}(0.5)$$

- Distribution

$$y \mapsto f_{Y|X}(y)$$

Fonctions de pertes:

- Quadratique:

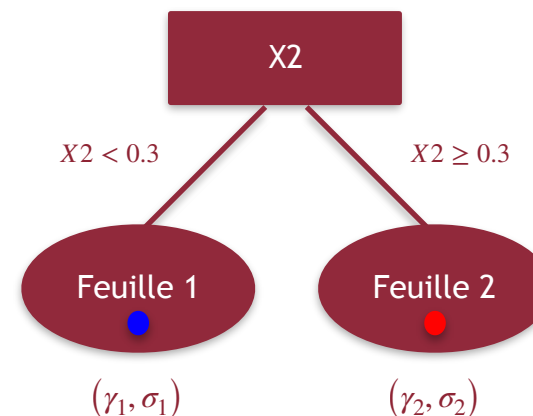
$$\phi(y, m(\mathbf{x})) = (y - m(\mathbf{x}))^2$$

- Absolue:

$$\phi(y, m(\mathbf{x})) = |y - m(\mathbf{x})|$$

- Vraisemblance

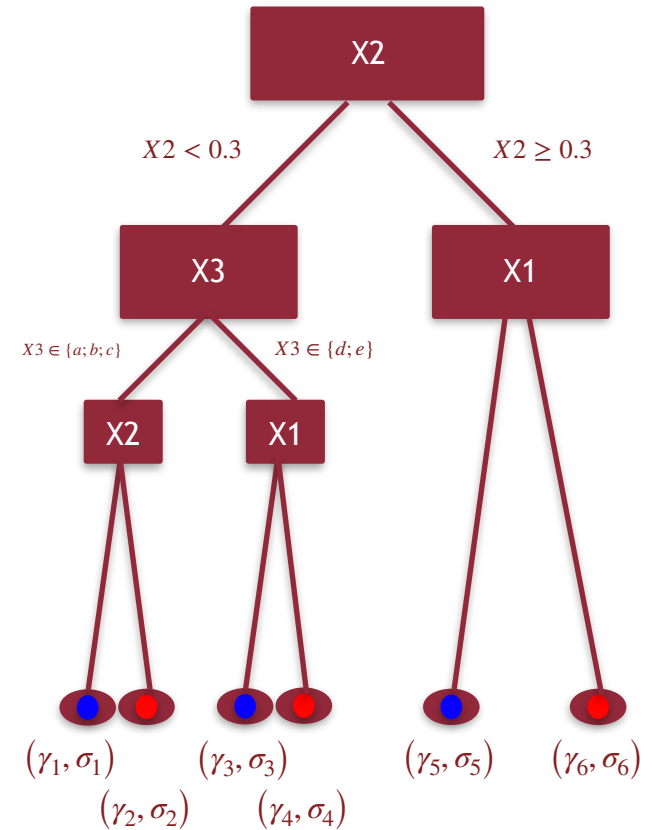
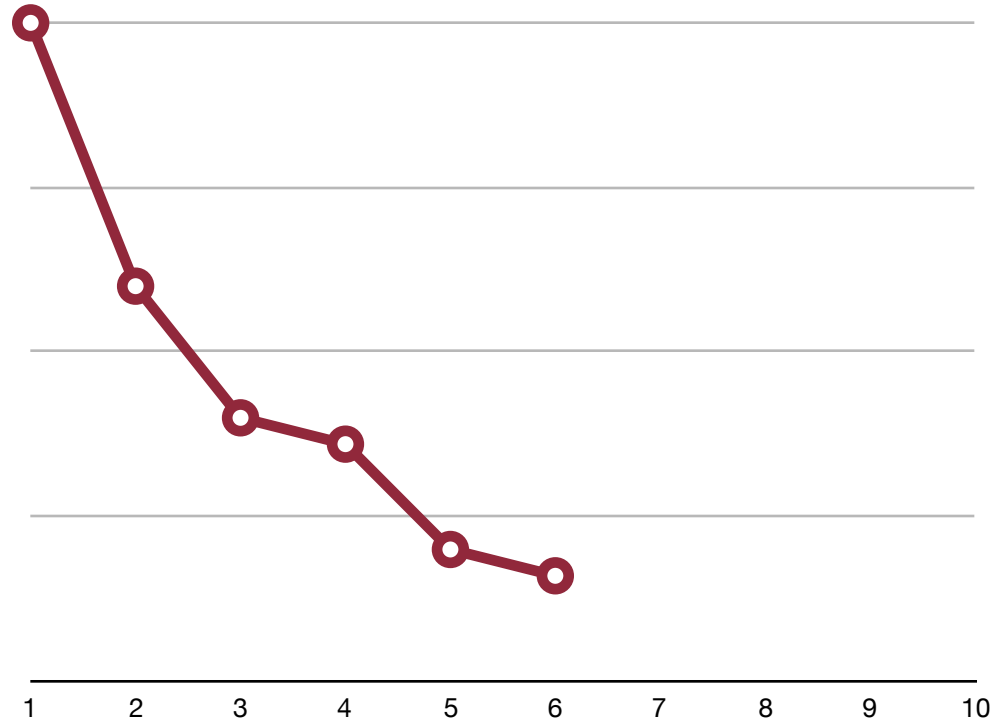
$$\phi(y, m(\mathbf{x})) = -\log f_{m(\mathbf{x})}(y)$$





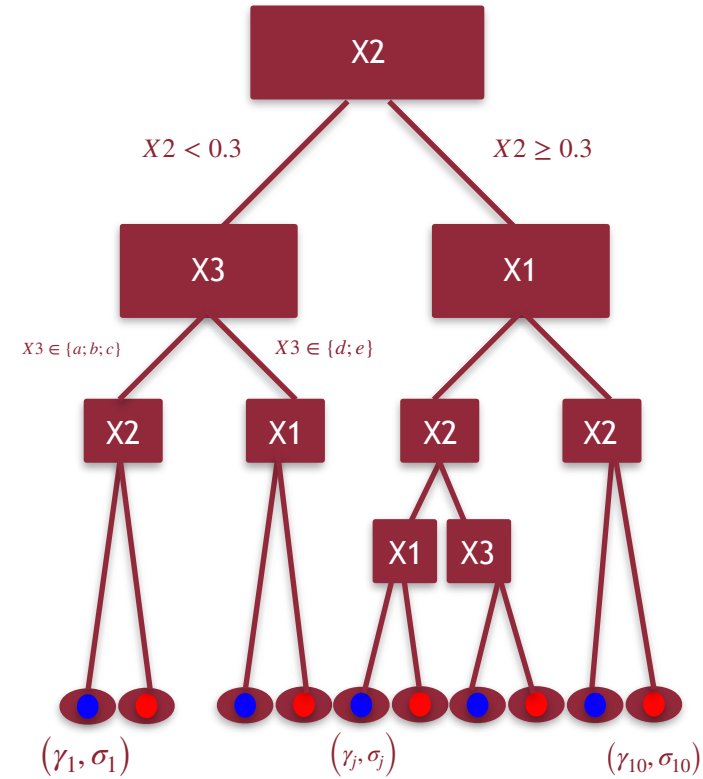
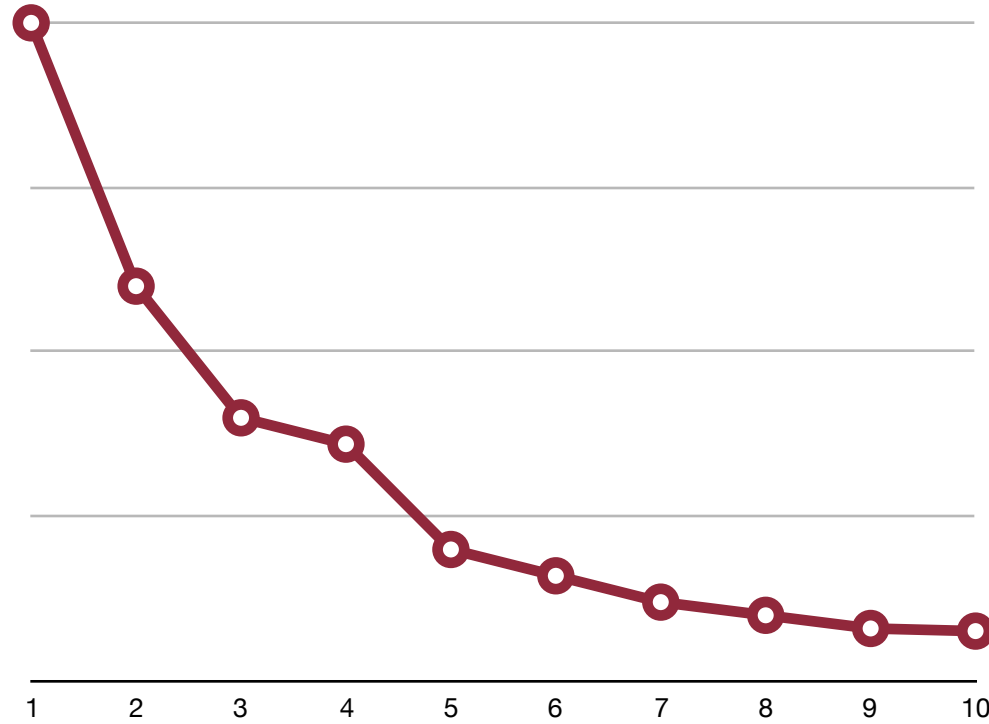


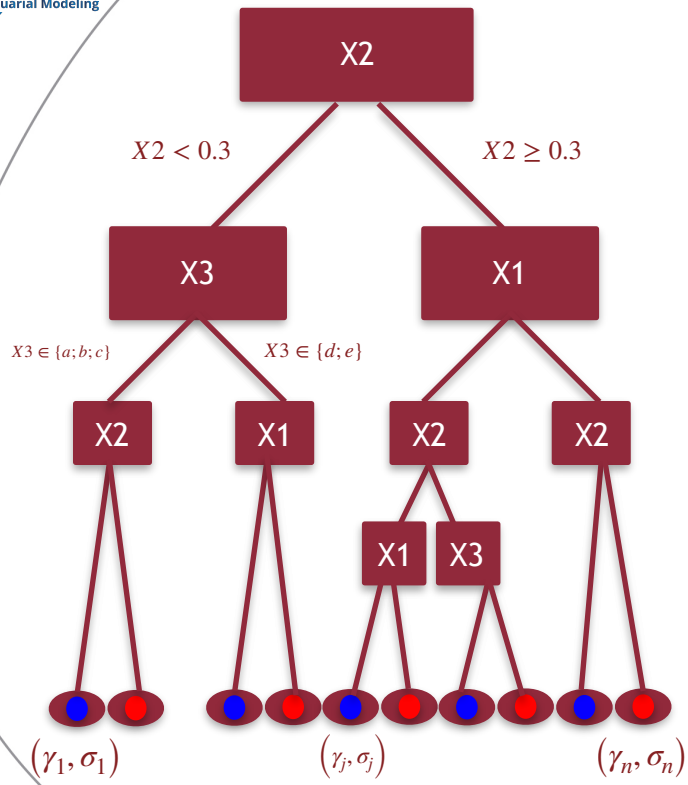
Réduction de l'erreur au fur et à mesure de l'accroissement de l'arbre



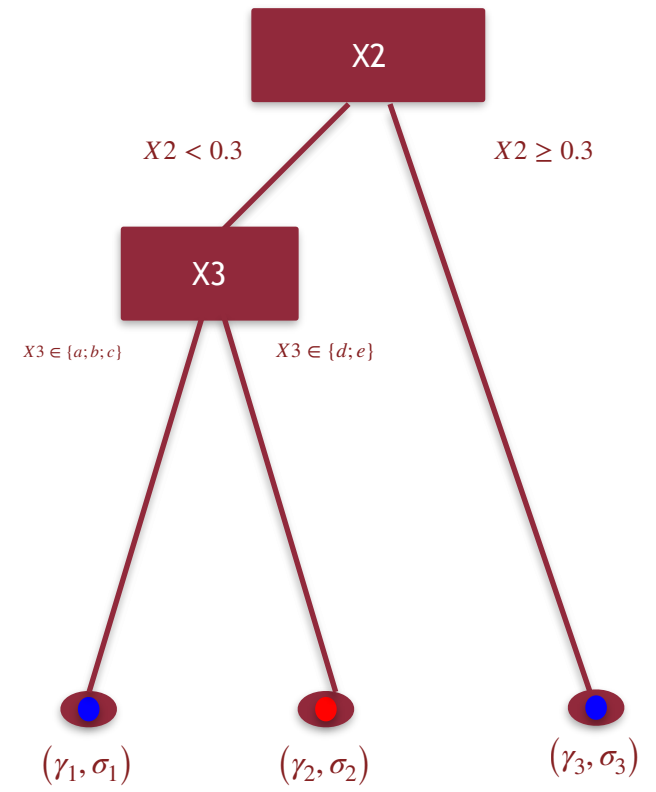
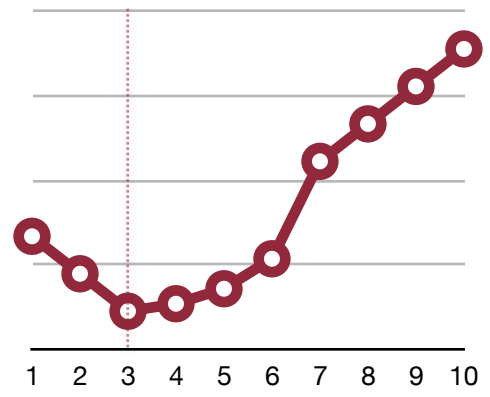
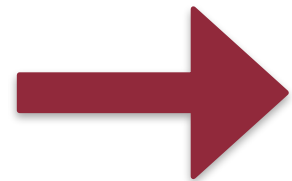


Réduction de l'erreur au fur et à mesure de l'accroissement de l'arbre





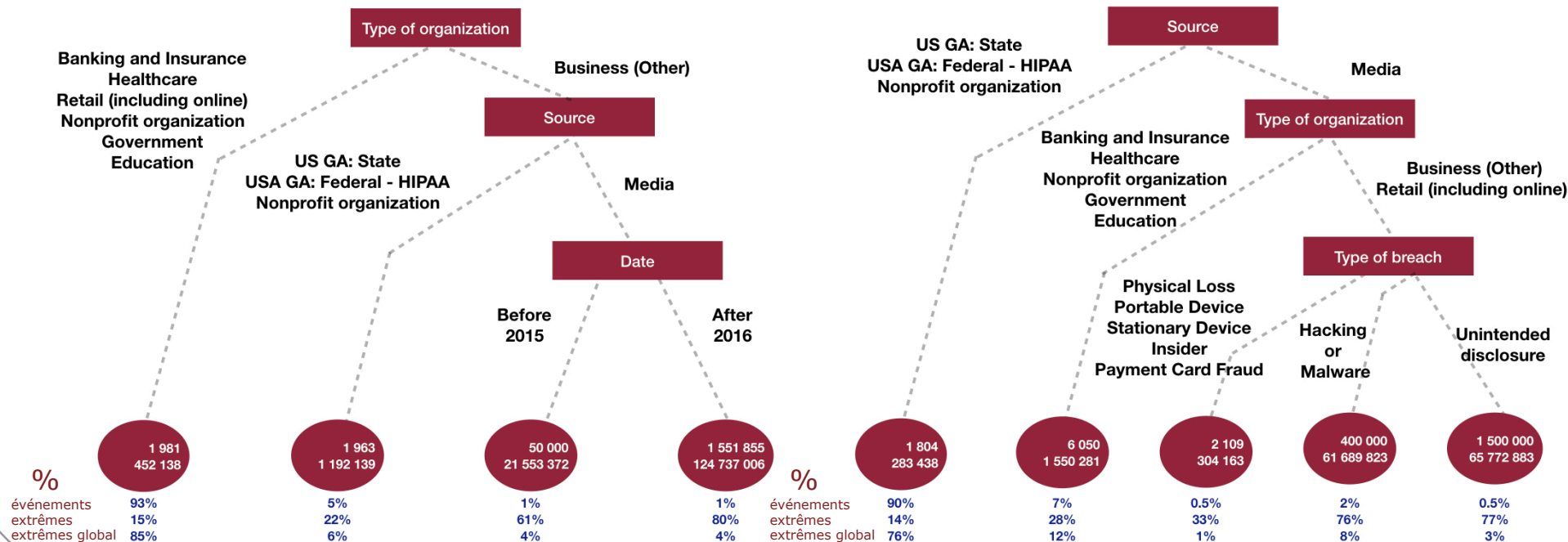
Validation croisée





Fonction de perte quadratique

Fonction de perte absolue

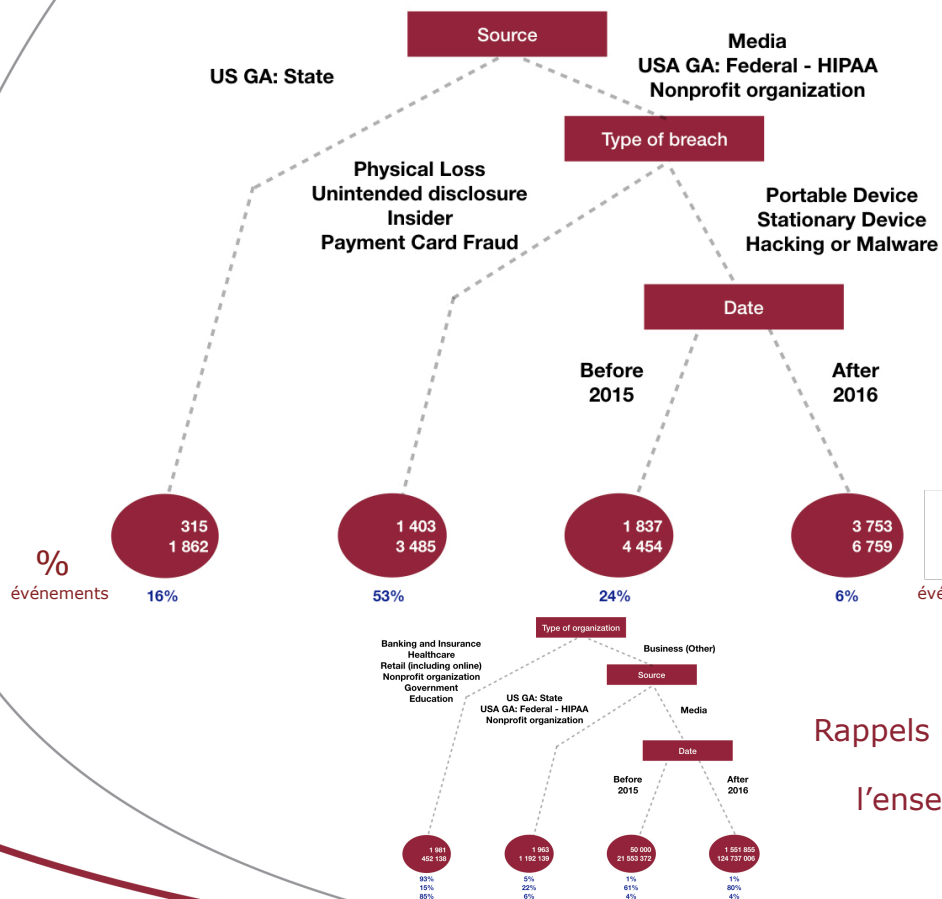


Médiane: 2 000  
Moyenne: 1 821 682

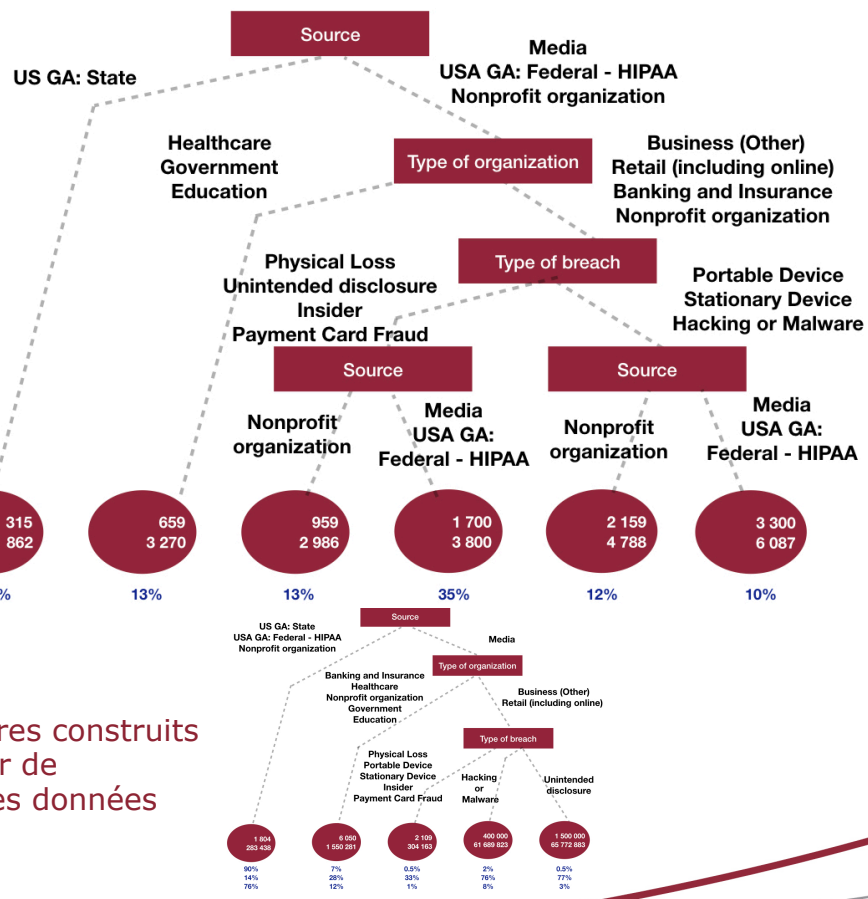


Fonction de perte quadratique

Fonction de perte absolue



Rappels des arbres construits à partir de l'ensemble des données

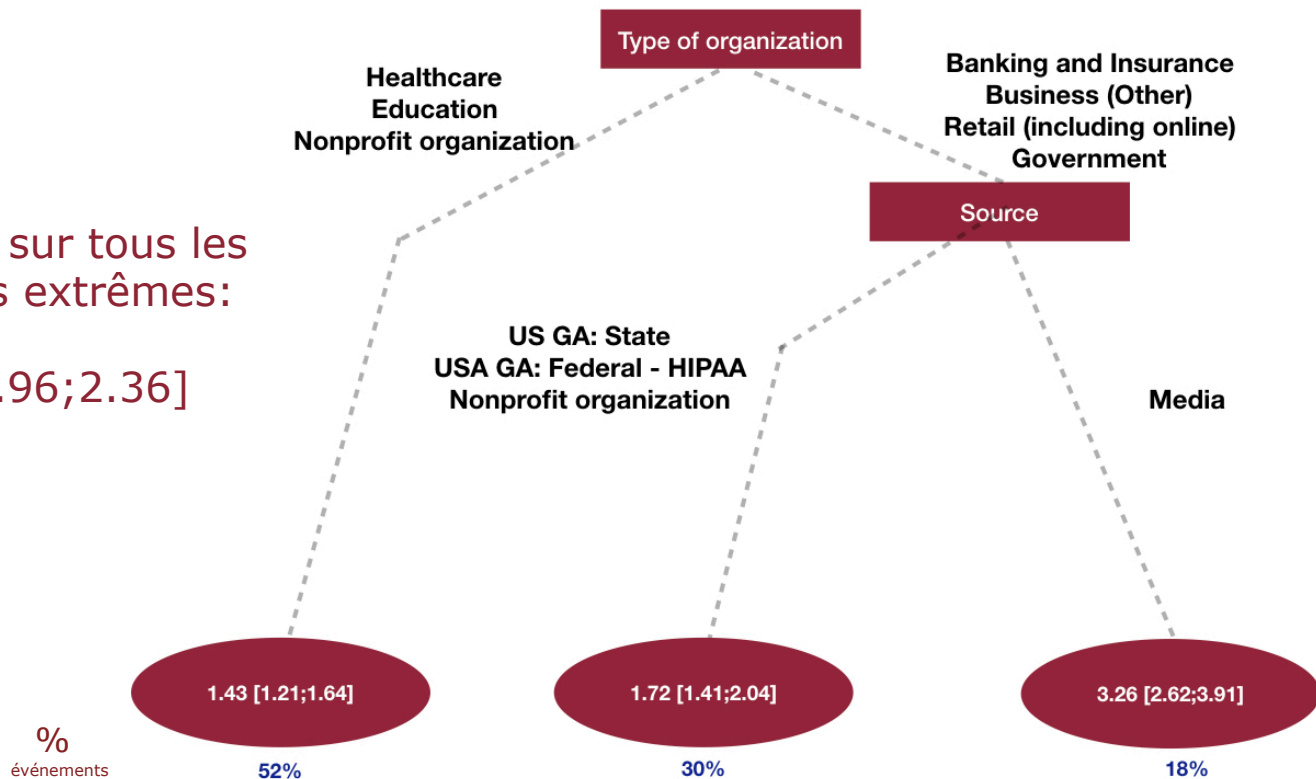




Fonction de perte vraisemblance  
à une loi de Pareto Généralisée

Ajustement sur tous les  
événements extrêmes:

2.16 [1.96;2.36]

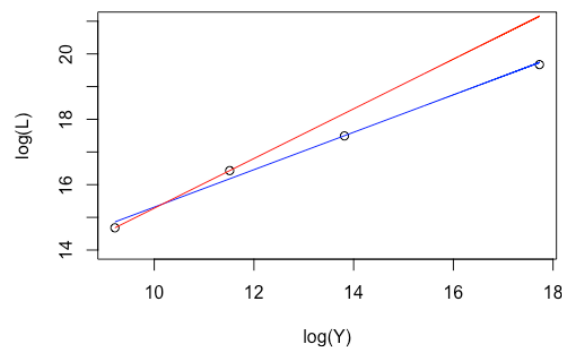




- Objectif: proposer un lien entre:
  - Y: le nombre de lignes d'une base révélée
  - L: le coût subit par l'entreprise victime
- Formule existente: Jacobs (2014):  $\log(L) = 7.68 + 0.76 \log(Y)$

	Moderate breaches		Mega breaches	
Number of records	10 000	100 000	1 000 000	50 000 000
Costs (in \$)	2 373 458	13 657 827	39 490 000	350 000 000
Costs per record (in \$)	237	137	39	7

- Formule proposée:  $\log(L) = 9.59 + 0.57 \log(Y)$





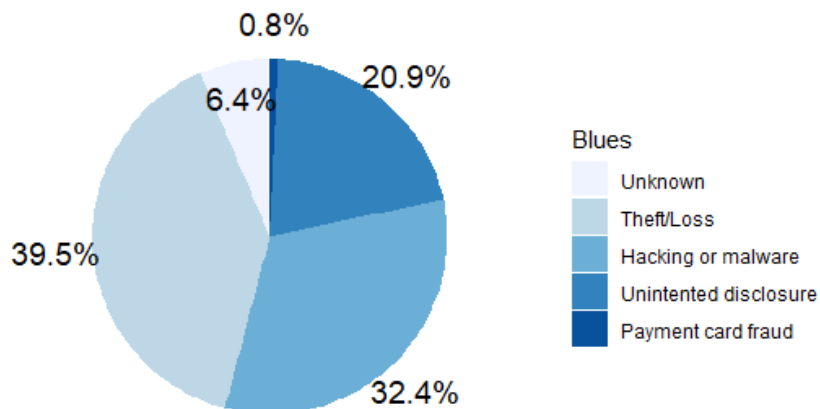


# Etude de la fréquence: calibrer le phénomène d'auto excitation

en collaboration avec  
Yannick Bessy-Roland (Axa)  
et  
Alexandre Boumezoued (Milliman)

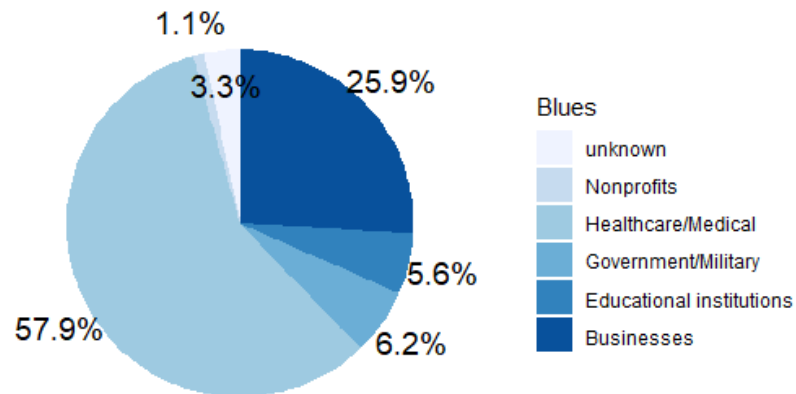
RESEARCH INITIATIVE  
Cyber Risk: Actuarial Modeling

### Types de brèches



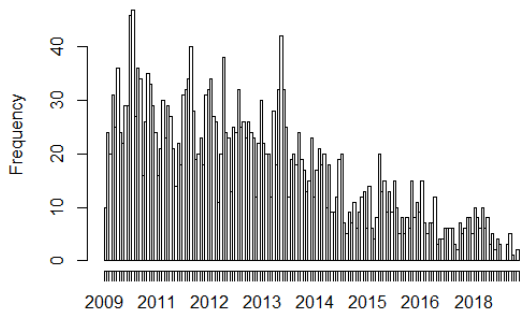
- Majorité de « Theft/Loss » et « Hacking/ Malware »
- 21% de fuites non intentionnelles

### Types d'organisations

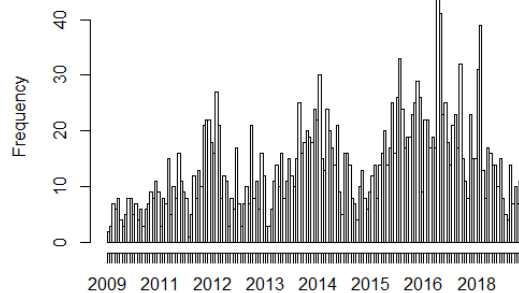


- Majorité de secteur médical
- Secteur Business bien représenté

Frequency of Theft/Loss

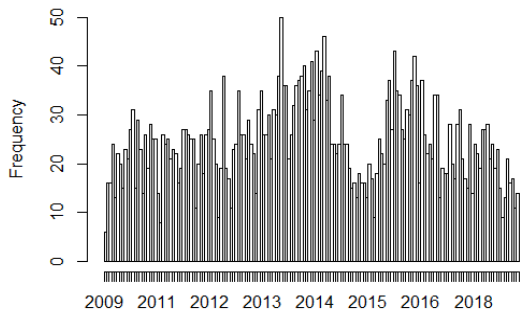


Frequency of Hacking/Malware

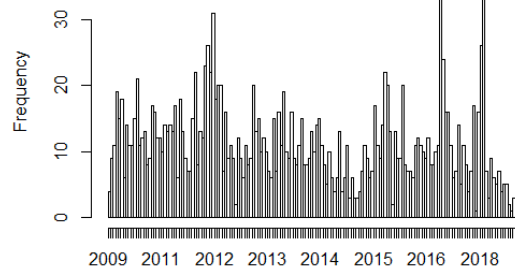


- clustering
- Tendance déterministe ou régime stochastique?

Frequency in Healthcare/Medical



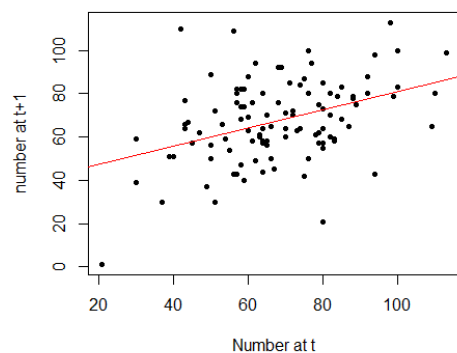
Frequency in Businesses



- clustering
- Pas de tendance claire

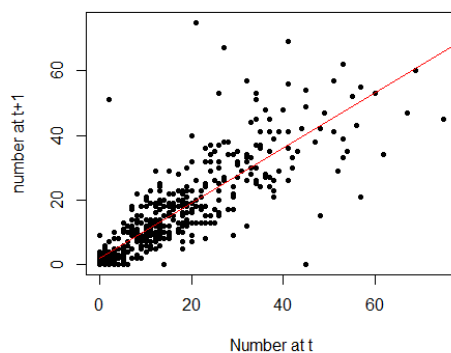
- Régression du nombre d'événements pendant le mois  $t + 1$  en fonction du nombre d'événements pendant le mois  $t$
- L'auto-corrélation augmente significativement si on se concentre sur les attaques de type/secteur **identiques**

Regression



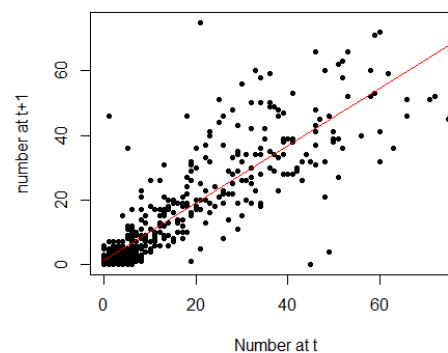
- R-squared : 0.154
- Intervalle de confiance (95%)  
[0.030, 0.278]

Regression per type of attack



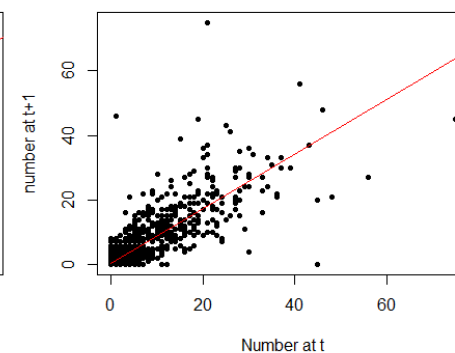
- R-squared : 0.726
- Intervalle de confiance (95%)  
[0.687, 0.766]

Regression per type of organisation



- R-squared : 0.780
- Intervalle de confiance (95%)  
[0.750, 0.810]

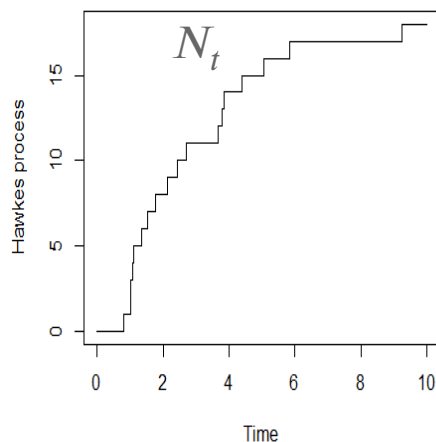
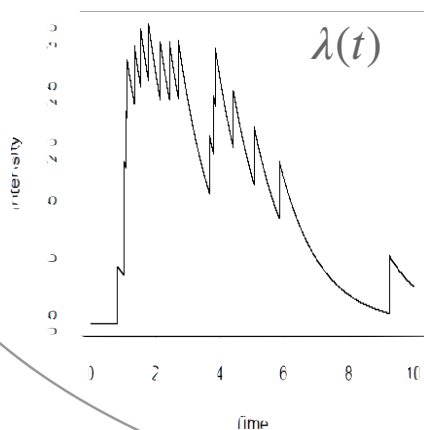
Regression per type of attack/types of organisation



- R-squared : 0.718
- Intervalle de confiance (95%)  
[0.702, 0.735]

- **Prise en compte de l'auto-corrélation**
  - Modèle de Cox : extension d'un processus de Poisson avec intensité stochastique → difficile de spécifier la dynamique de l'intensité.
  - Modèle de Hawkes : auto-excitant avec intensité stochastique, qui est entièrement spécifiée par le processus ponctuel.
- **Choix d'un modèle de Hawkes**
  - Auto-excitation: chaque événement augmente la probabilité d'occurrence d'un nouvel événement, au sein du même groupe (même secteur ou type d'attaques)
    - > Clustering
  - Inter-excitation: dans le cas de processus de Hawkes multi-dimensionnel, chaque événement au sein d'un groupe augmente la probabilité d'occurrence d'un nouvel événement au sein d'un autre groupe

- Processus de Hawkes univarié avec noyau exponentiel : processus de comptage  $N_t = \sum_{n \geq 1} 1_{T_n \leq t}$  de temps de sauts ( $T_n$ ) et d'intensité:
 
$$\lambda(t) = \mu(t) + \sum_{T_n < t} \alpha \exp(-\beta(t-T_n))$$
  - $\mu: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  : intensité de référence déterministe
  - La somme représente l'**impact des événements passés**, elle permet de capturer la propriété d'**auto-excitation**.



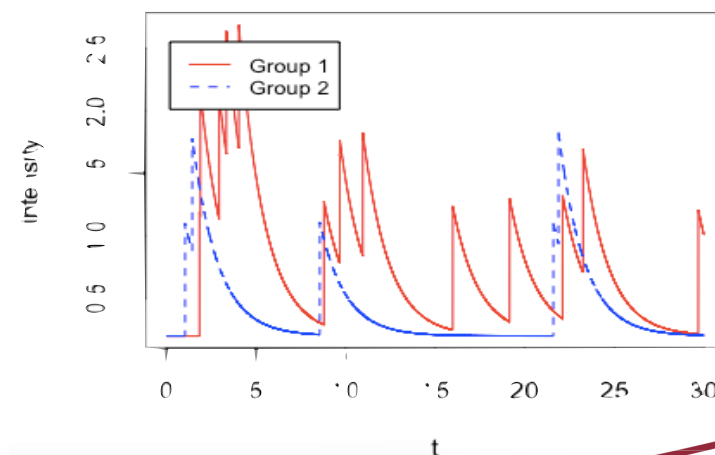
- Chaque saut représente un événement
- Phénomène de **clustering**
- Intensité décroît de façon **exponentielle** entre les sauts

- **Un processus de Hawkes multivarié** permet de modéliser des **interactions entre différents groupes**
- $(N_t^{(1)})_{t \geq 0}, \dots, (N_t^{(K)})_{t \geq 0}$ ,  $K$  processus de comptage de temps de sauts  $(T_n^{(1)})_{n \geq 1}, \dots, (T_n^{(K)})_{n \geq 1}$
- Avec noyau exponentiel: intensité  $\lambda_i(t)$  du processus  $(N_t^{(i)})_{t \geq 0}$

$$\lambda_i(t) = \mu_i(t) + \sum_{j=1}^K \sum_{T_n^{(j)} < t} \alpha_{i,j} \exp\{-\beta_{i,j}(t - T_n^{(j)})\}$$

- $\alpha_{i,j}, \beta_{i,j}$ : Impact du groupe  $j$  sur le groupe  $i$

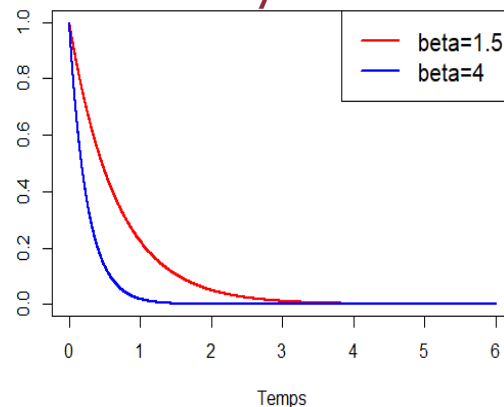
Intensity of Hawkes processes for 2 groups



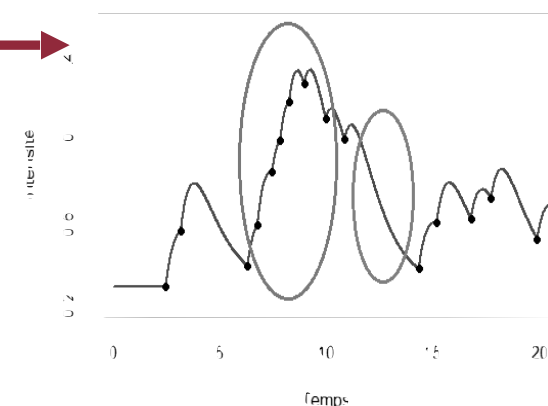
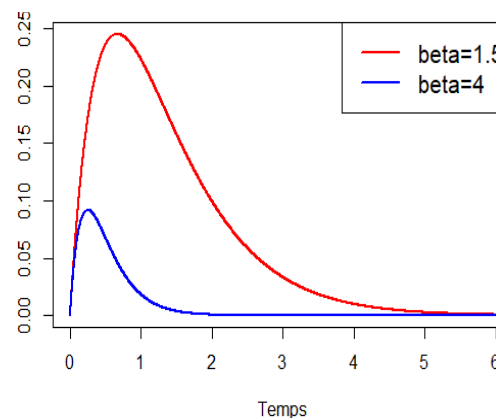
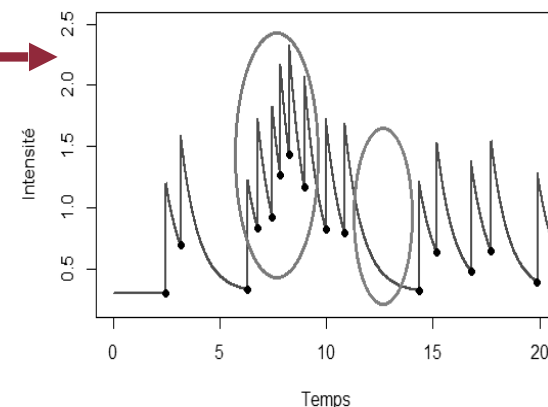


- **Noyau exponentiel « classique »**
- $\phi_{i,j}(s) = \alpha_{i,j} \exp(-\beta_{i,j}s)$
- Excitation instantanée
- Le processus intensité n'est pas Markov (pour dimension  $\geq 2$ )
- **Noyau avec délai:**
- $\phi_{i,j}(s) = \alpha_{i,j} s \exp(-\beta_i s)$
- Le processus intensité n'est pas Markov (même en dimension 1)

Noyaux



Processus Intensité



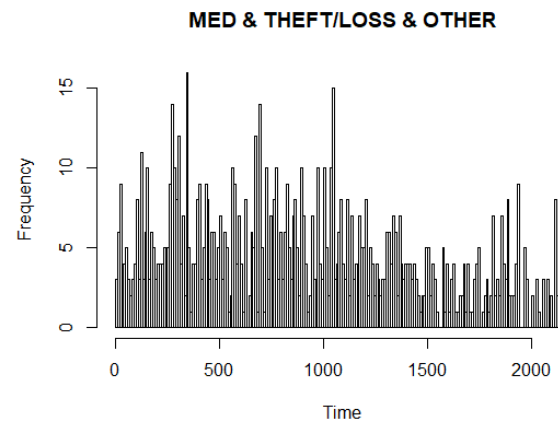
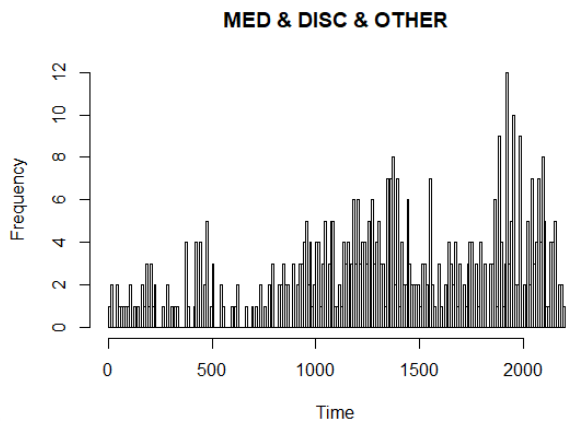
- **Variables:** type d'événement, secteur, localisation
- Total: **six groupes**

Group	Number of breaches
OTHER (1)	2046
MED & DISC & OTHER (2)	497
BUSINESSES & HACK & OTHER (3)	386
MED & HACK & OTHER (4)	472
MED & THEFT/LOSS & CALIFORNIA (5)	214
MED & THEFT/LOSS & OTHER (6)	943

- 3 noyaux considérés

	Kernel 1 (exp)	Kernel 2 (exp)	Kernel 3 (delay)
$\Phi_{i,j}(s)$	$\alpha_{i,j} \exp(-\beta_i s)$	$\alpha_{i,j} \exp(-\beta_{i,j} s)$	$\alpha_{i,j} s \exp(-\beta_i s)$
Nb paramètres	54	84	54

- **Intensité de référence** :  $\mu_i(t) = \mu_{0,i} + \gamma_i t$  pour modéliser la tendance



• **Vraisemblance**

	Kernel 1 (exp)	Kernel 2 (exp)	Kernel 3 (delay)
$\Phi_{i,j}(s)$	$\alpha_{i,j} \exp(-\beta_i s)$	$\alpha_{i,j} \exp(-\beta_{i,j} s)$	$\alpha_{i,j} s \exp(-\beta_i s)$
Nb parameters	54	84	54
-Likelihood (2011-2015)	6513	6172	<b>6153</b>
-Likelihood (2011-2016)	7639	7516	<b>7485</b>

• **Test d'adéquation (Kolmogorov-Smirnov)**

OTHER (1)	0.0503	0.0865	0.9060
MED & DISC & OTHER (2)	0.5546	0.1300	0.5173
BUSINESSES & HACK & OTHER (3)	0.5558	0.5966	0.3363
MED & HACK & OTHER (4)	0.0024	0.0361	0.0370
MED & THEFT/LOSS & California (5)	0.1146	0.5669	0.4246
MED & THEFT/LOSS & OTHER (6)	0.0733	0.6341	0.5379

Capture l'intensité de référence      même ordre de grandeur      Capture la tendance

	$\mu_0^{(i)}$	$\beta_i$	$\gamma_i$
OTHER (1)	0.87	5.39	-2.53e-04
MED & DISC & OTHER (2)	0.02	6.88	9.52e-05
Businesses & HACK & OTHER (3)	0.12	7.31	-3.56e-06
MED & HACK & OTHER (4)	0.02	5.75	9.65e-05
MED & Theft/Loss & CALIFORNIA (5)	0.05	5.96	-7.26e-06
MED & Theft/Loss & OTHER (6)	0.36	5.84	-1.07e-04

Table 7: Parameters  $(\mu_0^{(i)})_{1 \leq i \leq 6}$ ,  $(\beta_i)_{1 \leq i \leq 6}$  and  $(\gamma_i)_{1 \leq i \leq 6}$

	1	2	3	4	5	6
1	6.04	6.06	4.36	3.51	2.54	2.95
2	1.48	6.28	1.82	4.70	3.31	0.83
3	1.45	1.34	3.17	1.84	0.14	1.15
4	0.31	2.83	1.74	8.37	0.32	0.12
5	0.38	0.62	0.12	1.19	7.80	0.99
6	2.03	2.57	3.15	1.63	0.83	6.70

Table 8: Parameters  $(\alpha_{i,j})_{1 \leq i,j \leq 6}$

Capture les interactions

- Matrice des ratios entre excitation maximale et intensité de référence

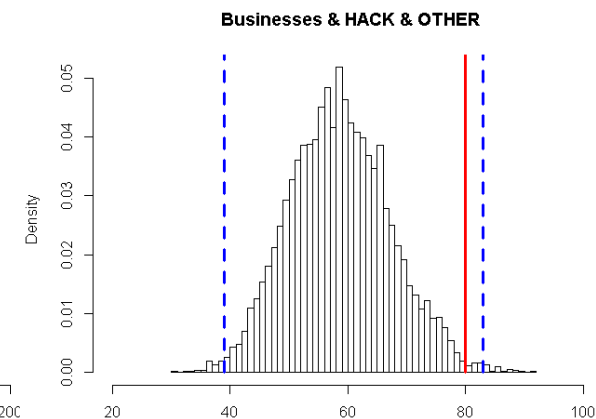
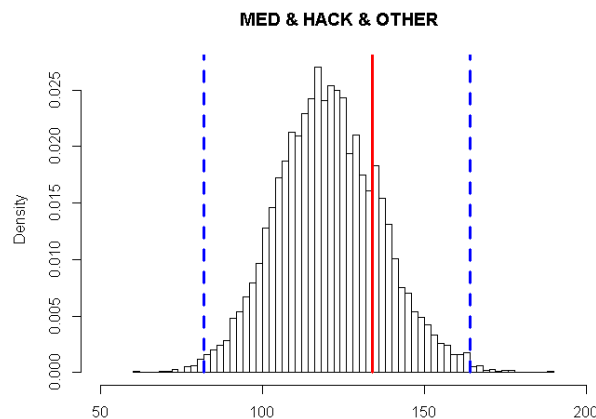
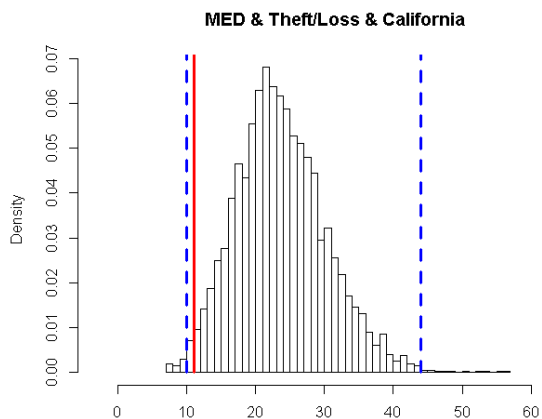
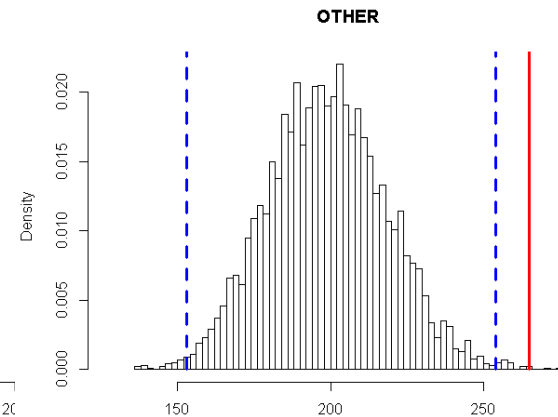
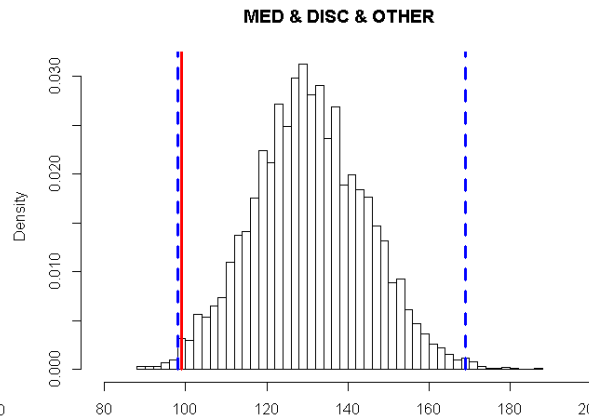
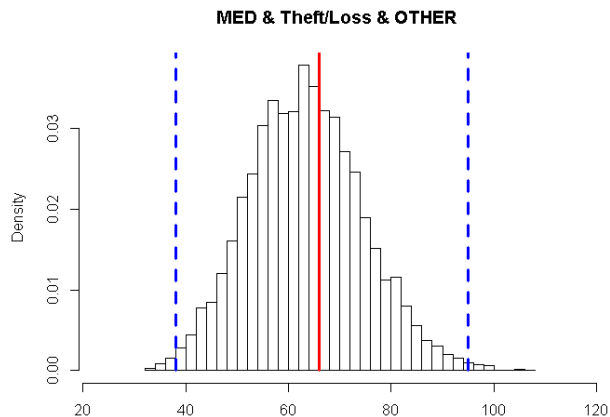
Forte auto-excitation entre groupes (2) et (4)

Excitation de (2) sur (5)

Forte auto-excitation pour ces groupes MED

	1	2	3	4	5	6
1	0.47	0.47	0.3	0.28	0.19	0.23
2	4	17	5	12.5	9	2
3	0.58	0.58	1.33	0.75	0.08	0.5
4	1	9	5.5	26.5	1	0.5
5	0.4	0.8	0.2	1.4	9.6	1.2
6	0.36	0.44	0.56	0.28	0.14	1.17

Table 10: Ratios Maximum excitation/basic intensity  $(\frac{\Gamma_{i,j}}{\mu_0^{(i)}})_{1 \leq i,j \leq 6}$



# Comprendre la diffusion des événements Cyber et adapter la réponse

## Le caractère systémique du risque cyber

- Catastrophe cyber : phénomène massif et rapide.
- Du fait de sa contagion, possibilité de démutualisation.
- Mutualisation : schématiquement, les mauvais résultats sur certains contrats sont absorbés par de bons résultats sur d'autres (mathématiquement : suppose que les risques contenus dans le portefeuille sont « **indépendants identiquement distribués** »)
- Exemple de cas de démutualisation en assurance (hors cyber) : catastrophes naturelles et portefeuilles déséquilibrés géographiquement.



## Un exemple de démutualisation

- Voici mon portefeuille d'assurance risques naturels :



## Cas du risque cyber

- La notion de proximité géographique n'est plus la même.
- On ne peut pas exclure une contagion massive au sein du portefeuille (**risque d'accumulation**).

## Modèles SIR

- Les modèles SIR (Susceptible  $s(t)$  - Infected  $i(t)$  - Recovered  $r(t)$ ) sont des modèles de base en épidémiologie humaine.

- Vision déterministe :

$$\frac{ds(t)}{dt} = -\beta s(t)i(t), \frac{di(t)}{dt} = \beta s(t)i(t) - \gamma i(t), \frac{dr(t)}{dt} = \gamma i(t).$$

- Des visions stochastiques existent.
- Ces modèles sont en général plutôt adaptés à des grandes populations (population nationale).
- Part du principe que les infectés  $i(t)$  rencontrent les susceptibles  $s(t)$  et que les nouvelles infections sont liées au nombre de rencontres entre ces deux populations.

## Vision portefeuille

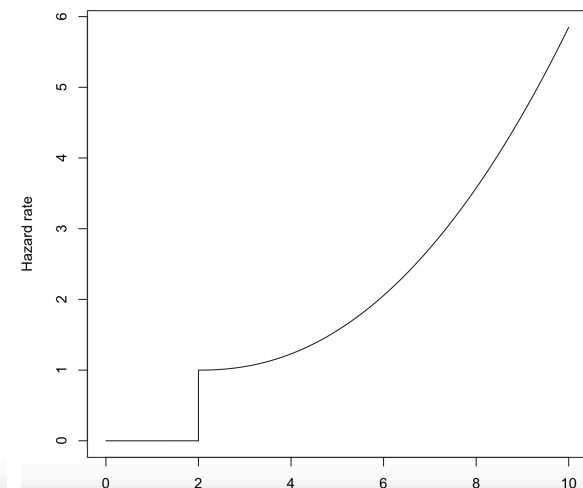
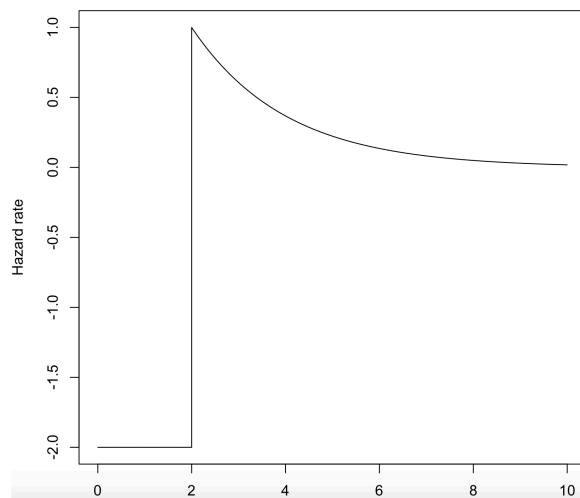
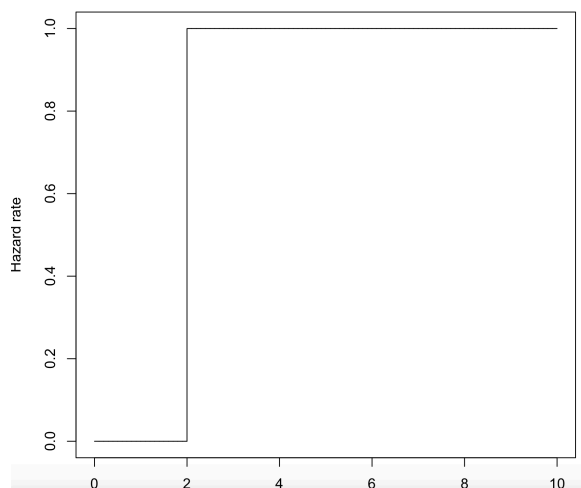
- La contagion vient moins de l'intérieur du portefeuille que du reste du monde (petite population).
- Modèle simple : les assurés sont classés en trois catégories.
- $S(t)$  = non infectés et vulnérables.
- $I(t)$  = infectés et des mesures d'assistance à court terme sont requises (par exemple pour redémarrer l'activité).
- $R(t)$  = l'activité a redémarré (même s'il peut rester des réparations à effectuer) ou les assurés ne sont plus vulnérables.
- **Risques concurrents** :  $S(t)$  désigne l'exposition, et pour passer dans la catégorie  $R$ , il y a « compétition » entre infection et prévention.

## Les quantités à modéliser

- $\lambda_I(t)dt$  = probabilité instantanée qu'un assuré vulnérable soit infecté à la date  $t$ .
- Typiquement,  $\lambda_I(t)$  dépend du nombre d'infectés dans la population nationale, qui peut être modélisée par un SIR (i.e.  $\lambda_I(t) = i(t)$ ).
- $\lambda_R(t)dt$  = probabilité instantanée qu'un assuré infecté soit « guéri ».
- Éventuellement,  $\lambda_R(t)$  devrait pouvoir dépendre de  $I(t)$ .
- $\lambda_P(t)dt$  = probabilité instantanée de voir la vulnérabilité corrigée avant infection.

## La réponse

- $\lambda_P$  modélise la capacité de réponse à l'événement.
- Trois exemples de modèles possibles :



## Le risque d'une saturation de la réponse

- Dans un problème de théorie du risque classique, chaque assuré a (en moyenne) le même coût  $C$ .
- Ce qui importe, c'est donc  $N(t)$  nombre total de victimes à la date  $t$ , et le coût est alors  $CN(t)$ .
- Ici, l'épisode a lieu sur un temps court. Si  $I(t)$  (nombre d'assurés à assister à la date  $t$ ) devient trop grand, le coût de réponse de l'assureur peut augmenter (voire l'assureur devient incapable d'assister ses assurés).
- Nécessité d'évaluer des probabilités du type  $\mathbb{P}(\sup_t I_t \geq s)$ , où  $s$  est un seuil qui correspond à une capacité de réponse.

## Quelques approximations

- À partir des dynamiques d'infection et de guérison, on peut approcher la distribution des processus  $S$ ,  $I$ , et  $R$ .

- Soit  $Z_t = \begin{pmatrix} N_t \\ R_t \end{pmatrix}$  où  $N_t$  est le nombre total d'infectés depuis le début de

l'épisode jusqu'à la date  $t$ . On a  $\frac{1}{N^{1/2}} \left\{ Z_t - N \begin{pmatrix} \nu(t) \\ \rho(t) \end{pmatrix} \right\}$  qui peut être

approché par un processus gaussien  $\mathcal{Z}$  (quand  $N$ , nombre d'assurés, est suffisamment grand) de structure de covariance:

$$\Sigma(t, h) = E \left[ (\mathcal{Z}_t^{(1)}, \mathcal{Z}_t^{(2)})^T (\mathcal{Z}_{t+h}^{(1)}, \mathcal{Z}_{t+h}^{(2)}) \right] = \begin{pmatrix} \nu(t)(1 - \nu(t+h)) & \phi_{Y,U}(t, h) - \nu(t)\rho(t+h) \\ \rho(t)(1 - \nu(t+h)) & \rho(t)(1 - \rho(t+h)) \end{pmatrix},$$

où

$$\nu(t) = \int_0^t S_C(u) f_T(u) du, \rho(t) = \int_0^t S_C(u) f_V(u) du, \phi_{Y,U}(t, h) = \int_0^t S_C(u) \{ F_U(t+h) - F_U(u) \} f_T(u) du.$$

## La nécessité d'une épidémiologie cyber

- Parmi les paramètres qui définissent la dynamique, il faut définir  $\lambda_I(t)$  (ou  $i(t)$ ) dans la logique du modèle SIR).
- Nécessite de développer des typologies réalistes d'infection.
- Permet ensuite de calibrer une réponse adaptée (temps de réaction à l'épisode).



# Conclusion

## **Les modèles présentés donnent des pistes sur comment nous pouvons modéliser différents types d'évènements Cyber:**

- une fréquence annuelle pour les sinistres attritionnels (modèle de Hawkes)
- une sévérité pour les sinistres attritionnels ou atypiques (modélisation de la sévérité par arbres de régression)
- un évènement d'accumulation de type catastrophe (modèle épidémiologique)

Comment faire le lien entre des modèles calibrés sur données de marché et le risque propre d'un portefeuille d'assurance ?

- risque de base
- approche de type crédibilité

En parallèle, étant donné le caractère évolutif du risque Cyber d'autres modèles continueront à voir le jour.

Merci de votre attention